

## **GENDEX-- CUT MODULE FOR CONSTRUCTING BLOCKED TREATMENT DESIGNS**

By

Walter T. Federer  
Department of Biometrics, Cornell University,

Nam-Ky Nguyen  
Designcomputing, and

Sophonie Nshinyabakobeje  
Department of Biometrics, Cornell University

### **ABSTRACT**

The CUT module of the Gendex toolkit may be used to block a factorial treatment design, a fractional factorial (or fractional replicate) treatment design, a response surface treatment design, or an  $n \times m$  array of data into  $b$  blocks. The blocking algorithm used makes the block effects orthogonal or near-orthogonal to the treatment effect parameters. The parameter vector consists of the block effects, the linear regression effects of the  $m$  variables (factors), the interactions of the  $m$  variables, and the quadratic regression effects of the  $m$  variables. An input design or  $n \times m$  array is required to start the program. In the DOS directory, the input design may be typed or pasted in notepad and saved. This is the design that is arranged into  $b$  blocks. Several examples are presented to illustrate the various aspects of the program in the CUT module.

**Keywords:** Fractional factorial, Response surface design, Linear effect, Quadratic effect, Interaction, Orthogonality, Main effect, Interaction.

BU-1591-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca New York 14853, January 2002

## INTRODUCTION

CUT is a program for blocking factorials, fractional factorials (FFDs), and response surface designs (RSDs). The algorithm described by Nguyen (1995) is used to implement the approach taken for constructing blocked designs. The  $u^{\text{th}}$  row of the extended design matrix  $\mathbf{X}$ ,  $\mathbf{W}$ , for  $n$  runs in  $b$  blocks of sizes  $n_i$ ,  $i = 1, 2, \dots, n$ , involving  $p - b$  parameters ( $m$  input variables and  $p - b - m$  derived variables) is  $(z_{u1}, \dots, z_{ub}, x_{u1}, \dots, x_{um}, \dots, x_{u(p-b)})$ . The block variable  $z_{uw}$ ,  $w = 1, \dots, b$ , is a dummy variable taking on the value 1 if  $u$  belongs to block  $w$  and 0 otherwise. The runs should be allocated to blocks such that block variables and  $x$ -variables are orthogonal. For an orthogonal blocked design, the inclusion of blocks does not affect the estimated regression coefficients of the  $x$ -variables and as such the only effect of blocking is reducing experimental error (Box and Hunter, 1957, page 228).

$\mathbf{W}$  is partitioned as  $[\mathbf{Z}|\mathbf{X}]$ . Let  $\mathbf{M} = \mathbf{W}'\mathbf{W}$ .  $\mathbf{M}$  can then be partitioned as  $\mathbf{M}_{11} = \mathbf{Z}'\mathbf{Z}$ ,  $\mathbf{M}_{12} = \mathbf{Z}'\mathbf{X}$ ,  $\mathbf{M}_{21} = \mathbf{X}'\mathbf{Z}$ , and  $\mathbf{M}_{22} = \mathbf{X}'\mathbf{X}$ .  $\mathbf{Z}'\mathbf{X}$  can be written as

$$\begin{pmatrix} s_{11} & s_{12} & \dots & s_{1(p-b)} \\ s_{21} & s_{22} & & s_{2(p-b)} \\ \vdots & & \vdots & \\ s_{b1} & s_{b2} & \dots & s_{b(p-b)} \end{pmatrix}$$

where  $s_{wj}$  is the contribution of block  $w$  to  $\sum_{u=1}^n x_{uj}$ , i. e.,  $s_{wj} = \sum_{u=1}^{n_w} x_{uj}$  ( $w = 1, \dots, b$ ;  $j = 1, \dots, p - b$ ). The condition of orthogonal blocking of a design has been shown to be  $s_{wj} - n_w/n s_j = 0$  where  $s_j = \sum_{w=1}^b s_{wj}$  ( $= \sum_{u=1}^n x_{uj}$ ) (See Equation 8 of Box and Hunter, 1957). Basically this condition states that the contribution of block  $w$  to each  $\sum_{u=1}^n x_{uj}$  must be proportional to the block size  $n_w$  (or must equal  $s_j/b$  in case of equal block sizes).

The approach to blocking a design with CUT is to find a suitable unblocked design and allocate the  $n$  runs of the design to  $b$  blocks such that  $f = \sum (s_{wj} - n_w/n s_j)^2$  is minimized.

## STEPS TO GENERATE A BLOCKED TREATMENT DESIGN

It is assumed that the CUT module is in a file called Gendex in the C:\ directory (See Federer *et al.*, 2001a, for instructions for installing the toolkit Gendex and creating the C:\Gendex> directory.). The various steps used to create a blocked design are:

1. An input file is required. An appropriate file may be obtained using the FFD module (Federer *et al.*, 2001b), the NOA module (Federer *et al.*, 2001c), the DEX module (Federer *et al.*, 2001d), or any appropriate design. This file is entered in NOTEPAD in the C:\Gendex> directory. The input file may be typed or pasted in notepad. The input design is saved as cut (which will appear as cut.txt) in the C:\Gendex directory.
2. Close NOTEPAD and in DOS and the C:\Gendex> directory, type cut, return.

3. A MENU will appear. Go to the C:\Gendex directory and select cut.txt. Click on OPEN.
4. A prompt, "cut.txt has n rows, (n runs), and m columns (m variables)", will appear giving the number of rows and columns of the input design cut.txt. Click on YES.
5. The next prompt is "choose the number of blocks". Select the required number of blocks and click OK.
6. The next prompt is "choose the model". The selections are Main effect model (The parameters are the linear regression effect of the m variables.), Interaction (The parameters are the linear effect of the m variables and the linear by linear interactions of the m variables.), Full model (This adds the quadratic effects of the m variables to the Interaction model parameters.), and Full model (main effect important) (This places more emphasis on lowering the standard errors of the main effect at the expense of other parameters). Select desired model and click OK.
7. The next prompt is "Enter a random seed". Enter a number or leave blank and click OK.
8. The last prompt is "Enter the number of tries". Enter a number or leave blank and click OK.
9. The design will appear on the screen. Click OK at bottom of screen to save the output. A description of the design just created and saved as CUT.HTM is given next along with DOS C:\Gendex> for creating another design.
10. To obtain a copy of CUT.HTM, go to WINDOW EXPLORER, C:\Gendex. At the bottom of the files in Gendex, highlight the CUT.HTM file and right click and select REFRESH.
11. Under FILE, OPEN CUT.HTM. If desired, a copy of the output may be printed. Under EDIT, click on SELECT ALL and then on COPY.
12. PASTE the output in the desired directory for editing or printing.

## OUTPUT FOR A BLOCKED TREATMENT DESIGN

The output for a created blocked design is

Note: the values of number of variables m, the number of parameters p, the number of runs n, and the number of blocks b.

Try #, gives the number of tries used to create the design.

Seed, gives the random seed selected or the one selected by the module.

# of iterations, gives the number of iterations for the various tries.

The value of f, the smaller the better the design, 0 for an orthogonal design. (f is defined above.)

Det is the determinant of  $\mathbf{X}'\mathbf{X}$ .

Std det.

BF

Factor levels, the block numbers are added to the input design.

The  $\mathbf{X}'\mathbf{X}$  matrix.

Inverse( $\mathbf{X}'\mathbf{X}$ ), which is the inverse of the  $\mathbf{X}'\mathbf{X}$  matrix.

## EXAMPLES

**Example 1.** The input treatment design for this example is a  $2^3$  complete factorial of  $n = 8$  combinations. The design is to be partitioned into two blocks,  $b = 2$ , such that all main effects and two-factor interactions are estimable. This means that the three-factor interaction ABC will be completely confounded with blocks. The steps used to create this blocked factorial are:

1. Go to MS-DOS Prompt and type CD\Gendex to change directory to C:\Gendex>.
2. Type NOTEPAD and return to open notepad.
3. In the open document either type the treatment design to be blocked or paste a design obtained from one of the other modules such as FFD. The design was typed in and saved as cut.txt in the Gendex directory.
4. Use the DOS command, C:\Gendex> CUT, return.
5. From the MENU, select CUT.TXT, open.
6. In response to the prompt, CUT.TXT has 8 rows and 3 columns, click on YES.
7. At the prompt, choose the number of blocks, 2 was selected.
8. At the prompt, choose the model, there are four options these are Main effect model, Interaction model, Full model, and Full model (main effects important). The option Full model was selected. For this example the Interaction and Full models will result in the same design.
9. At the prompt, Choose a random seed, either insert a number or leave blank and click on OK. This was left blank.
10. At the prompt, Choose the number of tries, either select a number or leave blank and click OK. This was left blank.
11. When the design appears on the screen, click on OK at the bottom of the screen and the design will be saved as CUT.HTM in the Gendex directory.
12. To obtain a Copy, go to WINDOWS EXPLORER and open the Gendex directory. Open the file CUT.HTM. Under EDIT, choose SELECT ALL and COPY.
13. Open the document where CUT.HTM is to be pasted. Under EDIT select PASTE.

Proceeding through the above steps, the following design was obtained.

CUT 2.3: partition a fractional factorial or response surface design into blocks  
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=8$ ,  $n=8$  and  $b=2$ .

```
try #           1
seed           1009815919200
# of iterations 2
f              0
det            4194304.0
std. det       0.25
BF             1
```

Factor levels (columns 2-4):

```
1   -1   -1   1
1   -1   1   -1
```

```

1  1  -1 -1
1  1   1  1
2 -1  -1 -1
2 -1   1  1
2  1  -1  1
2  1   1 -1

```

X'X

```

4  0  0  0  0  0  0  0
   4  0  0  0  0  0  0
     8  0  0  0  0  0
       8  0  0  0  0
        8  0  0  0
         8  0  0
          8  0
           8

```

inverse(X'X)

```

0.25  0  0  0  0  0  0  -0
      0.25  0  0  0  0  0  -0
           0.125  0  0  0  0  -0
               0.125  0  0  0  -0
                   0.125  0  0  -0
                       0.125  0  -0
                           0.125  -0
                               0.125

```

Note: CUT used 0.11 seconds.

Note: this software is licensed to AV Biometrics.

The **X** matrix for this treatment design is

```

1  0  -1  -1  1  1  -1  -1
1  0  -1  1  -1  -1  1  -1
1  0  1  -1  -1  -1  -1  1
1  0  1  1  1  1  1  1
0  1  -1  -1  -1  1  1  1
0  1  -1  1  1  -1  -1  1
0  1  1  -1  1  -1  1  -1
0  1  1  1  -1  1  -1  -1

```

For blocks, the presence or absence of a block is denoted by the dummy variable 1 or 0. The parameter vector  $\mathbf{B} = [b1 \ b2 \ A \ B \ C \ AB \ AC \ BC]$  where  $b1$  is block 1 effect and  $b2$  is block 2 effect.

**Example 2.** The input treatment design is a  $3^3$  factorial with  $n = 27$  combinations (runs). This design is to be blocked into  $b = 3$  blocks. The "Interaction model" is selected. Following the above steps, the output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks

(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=9$ ,  $n=27$  and  $b=3$ .

```
try #      80
seed      -2132268057
# of iterations  4
f          0
det        7.346640384E9
std. det    9.6341834E-4
BF         1
```

Factor levels (columns 2-4):

```
1  -1  -1  1
1  -1   0 -1
1  -1   1  1
1   0  -1  0
1   0   0 -1
1   0   1 -1
1   1  -1  0
1   1   0  0
1   1   1  1
2  -1  -1 -1
2  -1   0  1
2  -1   1 -1
2   0  -1  1
2   0   0  0
2   0   1  1
2   1  -1 -1
2   1   0  1
2   1   1 -1
3  -1  -1  0
3  -1   0  0
3  -1   1  0
3   0  -1 -1
3   0   0  1
3   0   1  0
3   1  -1  1
3   1   0 -1
3   1   1  0
```

$X'X$

```
9  0  0  0  0  0  0  0  0
   9  0  0  0  0  0  0  0
    9  0  0  0  0  0  0
     18 0  0  0  0  0
      18 0  0  0  0
       18 0  0  0
        12 0  0
         12 0
          12
```

$\text{inverse}(X'X)$

```
0.1111 0 0 0 0 0 0 0 -0
```

```

0.1111 0      0      0      0      0      0      -0
      0.1111 0      0      0      0      0      -0
            0.0556 0      0      0      0      -0
                  0.0556 0      0      0      -0
                        0.0556 0      0      -0
                              0.0833 0      -0
                                    0.0833 -0
                                          0.0833

```

Note: CUT used 0.99 seconds.

Note: this software is licensed to AV Biometrics.

Under Factor levels, the input design with the block numbers added in the first column is given. The parameter vector  $\mathbf{B} = [b_1 \ b_2 \ b_3 \ AL \ BL \ CL \ AL*BL \ AL*CL \ BL*CL]$  where VL is the linear effect of the variable V and VL\*UL is the interaction of the linear regression of variable V and the linear regression of the variable U. If the Main effects model had been selected, the interaction terms would not appear in the parameter vector  $\mathbf{B}$ .

**Example 3.** The input treatment design is a  $3^3$  factorial with 27 combinations. This design is to be blocked into  $b = 3$  blocks. The "Full model" is selected. The output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks

(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=12$ ,  $n=27$  and  $b=3$ .

```

try #      14
seed      1549233234
# of iterations  4
f          0
det        1.586874322944E12
std. det    1.0572492E-5
BF          1

```

Factor levels (columns 2-4):

```

1  -1  -1  1
1  -1   0 -1
1  -1   1  0
1   0  -1  0
1   0   0  1
1   0   1 -1
1   1  -1 -1
1   1   0  0
1   1   1  1
2  -1  -1  0
2  -1   0  0
2  -1   1  1
2   0  -1 -1
2   0   0 -1
2   0   1  0

```

```

2  1  -1  1
2  1   0  1
2  1   1 -1
3 -1 -1 -1
3 -1  0  1
3 -1  1 -1
3  0 -1  1
3  0  0  0
3  0  1  1
3  1 -1  0
3  1  0 -1
3  1  1  0

```

X'X

```

9  0  0  0  0  0  0  0  0  6  6  6
   9  0  0  0  0  0  0  0  0  6  6  6
      9  0  0  0  0  0  0  0  0  6  6  6
         18 0  0  0  0  0  0  0  0  0  0
            18 0  0  0  0  0  0  0  0  0
               18 0  0  0  0  0  0  0
                  12 0  0  0  0  0
                     12 0  0  0  0
                        12 0  0  0
                           18 12 12
                              18 12
                                 18

```

inverse(X'X)

```

0.3333 0.2222 0.2222 0  0  0  0  0  0  -0.1111 -0.1111 -0.1111
      0.3333 0.2222 0  0  0  0  0  0  0  -0.1111 -0.1111 -0.1111
          0.3333 0  0  0  0  0  0  0  0  -0.1111 -0.1111 -0.1111
              0.0556 0  0  0  0  0  0  0  0  0  0 -0
                  0.0556 0  0  0  0  0  0  0  0  0  0 -0
                      0.0556 0  0  0  0  0  0  0  0  0  0 -0
                          0.0833 0  0  0  0  0  0  0  0  0  0 -0
                              0.0833 0  0  0  0  0  0  0  0  0  0 -0
                                  0.0833 0  0  0  0  0  0  0  0  0  0 -0
                                      0.1667 0  0  0  0  0  0  0  0  0  0 -0
                                          0.1667 -0
                                              0.1667

```

Note: CUT used 0.66 seconds.

Note: this software is licensed to AV Biometrics.

In selecting the Full model over the Interaction model, the parameter vector is  $\mathbf{B} = [b_1 \ b_2 \ b_3 \ AL \ BL \ CL \ AL*BL \ AL*CL \ BL*CL \ AQ \ BQ \ CQ]$  where VQ is the quadratic regression effect for variable V (V = A, B, or C.) and are the terms added to the parameter vector for the Interaction model. The element under Factor levels for linear regression of variable V is squared to obtain the elements in the last three columns of the  $\mathbf{X}$  matrix used to obtain  $\mathbf{X}'\mathbf{X}$ .



**Example 4.** A  $3^3$  factorial set of  $n = 27$  combinations is used as the input design to illustrate the effect of selecting the Full model (main effects important). The output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks  
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=12$ ,  $n=27$  and  $b=3$ .

```
try #      26
seed      623735823
# of iterations  5
f         2*
det       1.5574877614080005E12
std. det   1.0376705E-5
BF        0.9979
```

Factor levels (columns 2-4):

```
1  -1  0  0
1  -1  0  1
1  -1  1  0
1   0 -1 -1
1   0 -1  0
1   0  0 -1
1   1 -1  1
1   1  1 -1
1   1  1  1
2  -1 -1 -1
2  -1  0 -1
2  -1  1  1
2   0 -1  1
2   0  0  1
2   0  1  0
2   1 -1  0
2   1  0 -1
2   1  1  0
3  -1 -1  0
3  -1 -1  1
3  -1  1 -1
3   0  0  0
3   0  1 -1
3   0  1  1
3   1 -1 -1
3   1  0  0
3   1  0  1
```

X'X

```
9  0  0  0  0  0  0  0  0  6  6  6
   9  0  0  0  0  0  0  1  6  6  6
     9  0  0  0  0  0 -1  6  6  6
       18 0  0  0  0  0  0  0  0  0
        18 0  0  0  0  0  0  0  0
          18 0  0  0  0  0  0  0
```

```

12  0  0  0  0  0
    12  0  0  0  0
        12  0  0  0
            18  12  12
                18  12
                    18

```

inverse(X'X)

```

0.3333 0.2222 0.2222 0 0 0 0 0 0 -0.1111 -0.1111 -0.1111
      0.3344 0.2212 0 0 0 0 0 -0.0094 -0.1111 -0.1111 -0.1111
        0.3344 0 0 0 0 0 0.0094 -0.1111 -0.1111 -0.1111
          0.0556 0 0 0 0 0 0 0 0 -0
            0.0556 0 0 0 0 0 0 0 0 -0
              0.0556 0 0 0 0 0 0 0 -0
                0.0833 0 0 0 0 0 0 -0
                  0.0833 0 0 0 0 -0
                    0.0849 0 0 -0
                      0.1667 0 -0
                        0.1667 -0
                          0.1667

```

Note: CUT used 0.71 seconds.

Note: this software is licensed to AV Biometrics.

The parameter vector for this example is the same as the previous one. The only change in the output from the Full model is that there is a "1" and a "-1" in the column for the parameter BL\*CL and the variance of this contrast is increased slightly.

**Example 5.** A  $3^3$  factorial set of  $n = 27$  combinations is used as the input design which is to be blocked into  $b = 9$  blocks. The Full model is selected and is to be compared with the output in the next example where the Full model (main effects important) is selected. The output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks  
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=18$ ,  $n=27$  and  $b=9$ .

```

try #      5
seed      1580273223
# of iterations  11
f         42
det       9.57854099865598E12
std. det   1.64722E-13
BF        0.8467

```

Factor levels (columns 2-4):

```

1  -1  -1  1
1   0   0 -1
1   1   0  0
2  -1   1  0
2   0  -1 -1

```

```

2   1   0   1
3  -1   0   1
3   0   1  -1
3   1  -1   0
4  -1   1  -1
4   0  -1   0
4   1   1   1
5  -1   0  -1
5   0  -1   1
5   1   1   0
6  -1   1   1
6   0   0   0
6   1  -1   1
7  -1  -1   0
7   0   1   1
7   1   0  -1
8  -1   0   0
8   0   1   0
8   1  -1  -1
9  -1  -1  -1
9   0   0   1
9   1   1  -1

```

X'X

```

3   0   0   0   0   0   0   0   0   0  -1   0   1  -1  -1   2   1   2
   3   0   0   0   0   0   0   0   0   0   0   0  -1   1   1   2   2   2
       3   0   0   0   0   0   0   0   0   0   0  -1  -1  -1   2   2   2
           3   0   0   0   0   0   0   0   1   0   0   2   0   2   3   2
               3   0   0   0   0   0   0   0   0   0   1   1  -1   2   2   2
                   3   0   0   0   0   0   0   0   0   2  -2   0   0   2   2   2
                       3   0   0   0   0   0   0   0   0   1  -1   1   2   2   2
                           3   0   0   0   0   0   0   0   0   0  -1  -1  -1   1   2   2   1
                               3   0   0   0   0   0   0   0   0   0  -1   2   0   0   2   2   3
                                   18   0   0   0   0   0   0   0   0   0   0   0   0   0   0
                                       18   0   0   0   0   0   0   0   0   0   0   0
                                           18   0   0   0   0   0   0   0   0   0
                                               12   0   0   0   0   0   0   0
                                                   12   0   0   0   0
                                                       12   0   0   0
                                                           18   12   12
                                                               18   12
                                                                   18

```

inverse(X'X)

```

0.5694 0.1512 0.199  0.1863 0.2244 0.1477 0.2171 0.1276 0.2921 0  0.0213 0.0069 -0.0685 0.0304 0.0414 -0.1111 -0.0473 -0.1385
      0.6279 0.2293 0.34  0.2428 0.2786 0.2269 0.2128 0.2593 0 -0.0105 -0.0047 0.0406 -0.0609 -0.037 -0.1111 -0.1426 -0.1189
        0.5962 0.2641 0.2281 0.2595 0.2202 0.201 0.2518 0 -0.0036 -0.0037 0.0329 0.0192 0.0311 -0.1111 -0.122 -0.1196
          0.7931 0.342 0.301  0.2705 0.2215 0.3521 0 -0.0337 -0.0016 -0.0063 -0.1105 -0.0033 -0.1111 -0.2122 -0.1329
            0.6433 0.2113 0.2451 0.1569 0.3477 0 -0.0065 0.0046 -0.0632 -0.0596 0.0376 -0.1111 -0.1307 -0.1429
              0.6687 0.2  0.21  0.1927 0 -0.0085 -0.0519 0.0951 -0.0229 -0.0058 -0.1111 -0.1367 -0.1082
                0.6111 0.1979 0.3269 0 -0.003 0.0069 -0.0568 0.0194 -0.0295 -0.1111 -0.12 -0.1326
                  0.5498 0.1853 0 -0.0052 0.0175 0.0442 0.022 -0.0396 -0.1111 -0.1268 -0.0504
                    0.7933 0 -0.0033 0.033 -0.1226 -0.0213 0.01 -0.1111 -0.1211 -0.2124
                      0.0556 0  0  0  0  0  0  0  0  0  0
                        0.0586 0.0005 -0.0035 0.0078 0.0025 0  0.0092 -0.0003
                          0.0641 -0.0149 0.0026 -0.001 0  0.0014 -0.0026

```

```

0.1451 -0.0011 -0.0106 -0    -0.0104 0.0278
0.1194 0.0008 -0    0.0235 0.0072
0.1013 0    0.0075 -0.0083
0.1667 -0    -0
0.1942 -0.0009
0.1937

```

Note: CUT used 0.88 seconds.

Note: this software is licensed to AV Biometrics.

The parameter vector  $\mathbf{B} = [b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6 \ b_7 \ b_8 \ b_9 \ AL \ BL \ CL \ AL*BL \ AL*CL \ BL*CL \ AQ \ BQ \ CQ]$ .

**Example 6.** A  $3^3$  factorial set of  $n = 27$  combinations in  $b = 9$  blocks is used to demonstrate the use of the Full model (main effects more important) versus the Full model. The output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks  
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=18$ ,  $n=27$  and  $b=9$ .

```

try #          2
seed          181884747
# of iterations 12
f            54*
det          5.355700839935989E12
std. det     9.210189E-14
BF           0.7937

```

Factor levels (columns 2-4):

```

1  -1  1  -1
1  0  0  1
1  1  -1  0
2  -1  -1  1
2  0  0  -1
2  1  1  0
3  -1  0  1
3  0  -1  0
3  1  1  -1
4  -1  0  0
4  0  -1  -1
4  1  1  1
5  -1  0  -1
5  0  1  0
5  1  -1  1
6  -1  -1  0
6  0  1  -1
6  1  0  1
7  -1  -1  -1
7  0  1  1
7  1  0  0
8  -1  1  1
8  0  0  0
8  1  -1  -1

```

```

9   -1   1   0
9   0   -1   1
9   1   0  -1

```

X'X

```

3   0   0   0   0   0   0   0   0   0   0   0  -2   1  -1   2   2   2
    3   0   0   0   0   0   0   0   0   0   0   0   2  -1  -1   2   2   2
        3   0   0   0   0   0   0   0   0   0   0   0   1  -2  -1   2   2   2
            3   0   0   0   0   0   0   0   0   0   0   0   1   1   2   2   2
                3   0   0   0   0   0   0   0   0   0   0  -1   2  -1   2   2   2
                    3   0   0   0   0   0   0   0   0   0   0   1   1  -1   2   2   2
                        3   0   0   0   0   0   0   0   0   0   0   1   1   2   2   2
                            3   0   0   0   0   0   0   0   0   0  -2  -2   2   2   2   2
                                3   0   0   0   0   0   0   0   0   0  -1  -1  -1   2   2   2
                                    18   0   0   0   0   0   0   0   0   0   0   0   0
                                        18   0   0   0   0   0   0   0   0   0   0   0
                                            18   0   0   0   0   0   0   0   0   0
                                                12   0   0   0   0   0   0   0
                                                    12   0   0   0   0   0
                                                        12   0   0   0
                                                            18   12   12
                                                                18   12
                                                                    18

```

inverse(X'X)

```

0.6667  0.1481  0.1667  0.1667  0.3148  0.2222  0.1667  0.2222  0.2593  0   0   0   0.1111  -0.0556  0.0556  -0.1111  -0.1111  -0.1111
0.6667  0.3148  0.2037  0.1667  0.2593  0.2037  0.1481  0.2222  0   0   0   -0.1111  0.0556  0.0556  -0.1111  -0.1111  -0.1111
0.6667  0.1667  0.1481  0.2222  0.1667  0.2222  0.2593  0   0   0   -0.0556  0.1111  0.0556  -0.1111  -0.1111  -0.1111
0.6667  0.2037  0.2222  0.3333  0.2222  0.1481  0   0   0   -0.0556  -0.0556  -0.1111  -0.1111  -0.1111  -0.1111
0.6667  0.2593  0.2037  0.1481  0.2222  0   0   0   0.0556  -0.1111  0.0556  -0.1111  -0.1111  -0.1111
0.6111  0.2222  0.1111  0.2037  0   0   0   0   -0.0556  -0.0556  0.0556  -0.1111  -0.1111  -0.1111
0.6667  0.2222  0.1481  0   0   0   0   0   -0.0556  -0.0556  -0.1111  -0.1111  -0.1111  -0.1111
0.7778  0.2593  0   0   0   0   0   0   0.1111  0.1111  -0.1111  -0.1111  -0.1111  -0.1111
0.6111  0   0   0   0.0556  0.0556  0.0556  -0.1111  -0.1111  -0.1111
0.0556  0   0   0   0   -0   0   0   0   0   0   -0
0.0556  0   0   0   0   -0   0   0   0   0   0   0   -0
0.0556  0   0   0   0   -0   0   0   0   0   0   0   -0
0.1667  -0   0   0   0   0   0   0   0   0   0   0
0.1667  0   0   0   0   0   0   0   0   0   0   0
0.1667  -0   -0   -0   -0   -0   -0   -0   -0   -0   -0
0.1667  -0   -0   -0   -0   -0   -0   -0   -0   -0   -0
0.1667  -0   -0   -0   -0   -0   -0   -0   -0   -0   -0
0.1667

```

Note: CUT used 0.66 seconds.

Note: this software is licensed to AV Biometrics.

From the  $\text{inverse}(X'X)$  matrix, we note that the variances of the VQ ( $V = A, B,$  or  $C$ ) parameters are lower for BQ and CQ than they were for the Full model. For the VL parameters, the variances for BL and CL parameters are lower than for the Full model selection. Thus the emphasis has been placed on estimating main effects. The variances for the VL\*UL interactions are higher than for the Full model selection to off-set the reduction in variances for main effect parameters.

**Example 7.** A  $4^2$  factorial with  $n = 16$  combinations in  $b = 4$  blocks with the Full model selection is used to illustrate the parameter vector. Note that even though a cubic regression is estimable for four levels of a variable, CUT does not include more than linear and quadratic regressions. The output obtained was:

CUT 2.3: partition a fractional factorial or response surface design into blocks  
(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=2$ ,  $p=9$ ,  $n=16$  and  $b=4$ .

```
try #      1
seed      1010068637550
# of iterations  4
f         32
det       4.0420742529024E13
std. det   588.1992
BF        0.9879
```

Factor levels (columns 2-3):

```
1  -1  1
1  -1  3
1  -3 -3
1   3 -1
2  -3  1
2   1 -1
2   1 -3
2   3  3
3  -1 -1
3  -3 -1
3   1  3
3   3 -3
4  -1 -3
4  -3  3
4   1  1
4   3  1
```

$X'X$

```
4  0  0  0 -2  0  2  20  20
   4  0  0  2  0  2  20  20
     4  0  0 -2 -2  20  20
       4  0  2 -2  20  20
         80  0  0  0  0
          80  0  0  0
            400  0  0
              656  400
                656
```

$\text{inverse}(X'X)$

```
0.4491 0.1927 0.1947 0.1947 0.0064 0 -0.0013 -0.0195 -0.0195
      0.4491 0.1947 0.1947 -0.0064 0 -0.0013 -0.0195 -0.0195
          0.4491 0.1927 0 0.0064 0.0013 -0.0195 -0.0195
              0.4491 0 -0.0064 0.0013 -0.0195 -0.0195
```

```

0.0128 0      0      0      -0
      0.0128 0      0      -0
              0.0025 0      -0
                      0.0039 -0
                          0.0039

```

Note: CUT used 0.28 seconds.

Note: this software is licensed to AV Biometrics.

The parameter vector is  $\mathbf{B} = [b1 \ b2 \ b3 \ b4 \ AL \ BL \ AL*BL \ AQ \ BQ]$ .

**Example 8.** The CUT module may be used to group an  $m$  by  $n$  array of numbers into  $b$  blocks such that the block totals are as nearly equal as possible. Let the input file be an array of  $n = 10$  rows by  $m = 3$  columns composed of 10 blends of three variables. It is desired to group the rows in such a manner as to have the block totals as nearly equal as possible. Suppose that  $b = 4$  groups of sizes 2, 2, 3, and 3 are desired of the 10 blends. Type in the array of  $mn = 30$  numbers in notepad and save as cut (file name will be cut.txt). In DOS type cut, return, and open cut.txt file in the C:\Gendex directory. The 30 numbers under factor levels constitute the input file. The following blocked array of the  $n = 10$  rows was obtained:

CUT 2.3: partition a fractional factorial or response surface design into blocks

(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=7$ ,  $n=10$  and  $b=4$ .

```

try #      1
seed      1010089244550
# of iterations 3
f         94
det       0.0
std. det  0.0

```

Factor levels (columns 2-4):

```

1   20   10   10
1   30   40   40
2   15   15   15
2   30   30   30
3   10   10   10
3   25   25   25
3   40   30   30
4   10   10   10
4   20   20   20
4   40   40   40

```

X'X

```

2   0   0   0   50   50   50
    2   0   0   45   45   45
        3   0   75   65   65
            3   70   70   70
                6850  6550  6550

```

6550 6550  
6550

Note: CUT used 0.17 seconds.

Note: this software is licensed to AV Biometrics.

Note that no inverse of  $\mathbf{X}'\mathbf{X}$  was obtained. Note that block totals for each of the  $m = 3$  variables appear in the upper right hand corner of the  $\mathbf{X}'\mathbf{X}$  matrix. Also note that the block totals for variables B and C are identical, meaning that the linear regressions for B and C are completely confounded and therefore the inverse does not exist.

**Example 9.** A 20/27 fraction of a  $3^3$  factorial in  $n = 20$  runs blocked into  $b = 5$  blocks with the Full model selection is used to illustrate blocking for a fractional replicate input design. The design was obtained using the FFD module. It was then copied and pasted in notepad and saved as cut. The cut.txt file was opened to obtain the following output:

CUT 2.3: partition a fractional factorial or response surface design into blocks

(C) 2001 Design Computing (URL: <http://designcomputing.hypermart.net/gendex>)

Note: design for  $m=3$ ,  $p=14$ ,  $n=20$  and  $b=5$ .

```
try #          3
seed          -1906243450
# of iterations 3
f             21.6
det           3.8843449344E10
std. det      2.370816E-8
BF            0.9078
```

Factor levels (columns 2-4):

```
1  -1  1  -1
1   0 -1  0
1   0  0  1
1   1  0  0
2  -1  0  0
2   0  0  1
2   0  1  0
2   1 -1 -1
3  -1  0  0
3   0 -1  0
3   0  0 -1
3   1  1  1
4  -1 -1 -1
4  -1  1  1
4   1 -1  1
4   1  1 -1
5  -1 -1  1
5   0  0 -1
5   0  1  0
5   1  0  0
```

$\mathbf{X}'\mathbf{X}$



```

4  0  0  0  0  0  0  -0  -1  1  -1  2  2  2
   4  0  0  0  0  0  -0  -1  -1  1  2  2  2
     4  0  0  0  0  -0  1  1  1  2  2  2
       4  0  0  0  -0  0  0  -0  4  4  4
         4  0  0  -0  1  -1  -1  2  2  2
           12 0  0  0  0  0  0  0  0  0
            12 0  0  0  0  0  0  0  0
              12 0  0  0  0  0  0  0
                8 0  0  0  0  0  0
                  8 0  0  0  0
                    8 0  0  0
                     12 8  8
                      12 8
                       12

```

inverse(X'X)

```

0.4643 0.1786 0.1786 0.375 0.1786 0  -0  0  0.0357 -0.0357 0.0357 -0.125 -0.125 -0.125
      0.4643 0.1786 0.375 0.1786 0  -0  0  0.0357 0.0357 -0.0357 -0.125 -0.125 -0.125
        0.4643 0.375 0.1786 0  -0  0  -0.0357 -0.0357 -0.0357 -0.125 -0.125 -0.125
          0.375 0  -0  0  -0  0  0  0  -0.25 -0.25 -0.25
            0.4643 0  -0  0  -0.0357 0.0357 0.0357 -0.125 -0.125 -0.125
              0.0833 0  0  0  0  0  0  -0  -0  -0
                0.0833 -0  0  -0  0  0  0  0  0  0
                  0.0833 -0  0  0  0  -0  -0  -0  -0
                    0.1429 -0  0  0  0  0  0
                      0.1429 -0  -0  -0  -0
                        0.1429 -0  -0  -0
                          0.25 0  0
                            0.25 0
                              0.25

```

Note: CUT used 0.55 seconds.

Note: this software is licensed to AV Biometrics.

The parameter vector  $\mathbf{B} = [b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ AL \ BL \ CL \ AL*BL \ AL*CL \ BL*CL \ AQ \ BQ \ CQ]$ .

## COMMENTS

It may be that CUT could be utilized in data mining to ascertain patterns in the rows of an array. It appears that further investigation is required to determine if the Gendex modules could be used to locate patterns in a data set. The CUT module is useful for blocking in many situations such as in industrial and pharmaceutical investigations. Often only the number of runs in a block can be performed at one time. The effects confounded with blocks are often considered to be negligible or non-existent.

## LITERATURE CITED

Box, G. E. P. and J. S. Hunter (1957). Multifactor response experimental designs for exploring response surfaces. *Annals of Mathematical Statistics* 8:195-241.

Federer, W. T., B. Gross, N-K. Nguyen, and S. Nshinyabakobeje (2001a). GENDEX-2001: Getting started and the BIB module. BU-1572-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca, New York 14853, September.

Federer, W. T., N-K. Nguyen, and S. Nshinyabakobeje (2001b). GENDEX: FFD module for constructing orthogonal or near-orthogonal fractional replicate and response surface treatment designs. BU-1574-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca, New York 14853, November.

Federer, W. T., N-K. Nguyen, and S. Nshinyabakobeje (2001c). GENDEX: NOA module for constructing orthogonal or near-orthogonal arrays and main effect plans. BU-1575-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca, New York 14853, November.

Federer, W. T., N-K. Nguyen, and S. Nshinyabakobeje (2001d). GENDEX: DEX module for constructing D-optimal treatment designs. BU-1576-M in the Technical Report Series of the Department of Biometrics, Cornell University, Ithaca, New York 14853, November.

Nguyen, N-K. (1995). An algorithmic approach to blocking response surface designs and other designs. CSIRO-IAPP Biometrics Unit Technical Report 95/21T.